Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 1812–1814, March 1991
Psychology

The speed of attentional shifts in the visual field

(psychophysics/human information processing/vision/attention)

JUKKA SAARINEN* AND BELA JULESZ†

Laboratory of Vision Research, Rutgers University, New Brunswick, NJ 08903

Contributed by Bela Julesz, November 30, 1990

ABSTRACT **The scanning speed of focal visual attention was measured directly by flashing a sequence of two, three, or four numerals one by one at random retinal positions and at a distance from each other to avoid interference between the numerals. Each numeral was followed by a mask pattern so the observers had to move their focal attention in the visual field in synchrony and at the same speed as the presentation rate of the numerals in order to recognize every numeral in the stimulus sequence. Observers could recognize the numerals orders of magnitude above the theoretical chance level of performance even when the presentation rate was as fast as 33 ms per numeral. However, the temporal order of the numerals was perceived rather poorly at the fast presentation rates and for the sequences of four numerals.**

That we can inspect element-by-element afterimages burned on our retina has already been reported by von Helmholtz (1), thus showing that shifts of focal attention are independent of eye movements. After a century of dormancy, there is new interest in focal attention. Several studies (2–8) have revealed an attentive mode of visual processing that is spatially limited. When focal attention is shifted to particular spatial position, stimuli in this location are processed more efficiently (i.e., faster and more accurately) than stimuli in other positions in the visual field (2, 3). It also seems that the *recognition* of stimuli, even a single stimulus feature (e.g., color or orientation), requires focal attention (6).

A much-studied aspect of attentive vision has been the speed of the attentional shifts in the visual field. However, there is no agreement what this speed is: the estimates have ranged from 50 to 300 ms per item (4, 5, 9–11). In this study, we investigated the speed of attentional shifts using the paradigm of *serial visual presentation*. Our paradigm resembles that of Sperling and his collaborators (9, 10), but instead of presenting alpha-numeric characters in the *same* window—that might slow down the speed of attentional shifts—we presented our targets at different retinal positions and far from each other in successive steps in order to reduce interference between the numerals. Furthermore, we did not use a cue to point at the selected position in the visual field, or in the temporal sequence, because the disengagement of attention from the cued location seems to be a separate mental operation that requires time (12).

In our procedure, two, three, or four numerals were flashed one by one in random positions of the visual field, and each numeral was followed by a mask pattern (Fig. 1). We assume that the observer had to move his focal attention in the visual field at the same speed as the presentation rate of numerals and in synchrony with them in order to recognize every numeral in the sequence. Hence, in this study, we used for didactic reasons the metaphor of the single "searchlight of attention," instead of some other metaphors—for example, the metaphor of "the limited resources of attention," which are allocated to different parts of the visual field. It has to be emphasized that our experimental results can be interpreted in several ways—i.e., we do not take sides whether our results represent "the speed of the searchlight of attention" or "the spatio-temporal allocation of finite attentional resources."

METHODS

The numerals and the mask patterns were each presented for 33 ms with a blank interval of variable duration between them (Fig. 1). The presentation rate of the numerals was varied by changing the length (0, 33, or 67 ms) of this blank interval. The 33-ms duration and its multiples were dictated by the non-interlaced frame frequency (60 Hz) of the Macintosh SE/30 computer and SuperMac 19-inch color Trinitron monitor, which were used to generate and to display the stimulus patterns. Because each numeral in the sequence was masked simultaneously as the next numeral appeared on the screen, the presentation rate of the numerals was defined to be the interval between the stimulus (numeral) and mask onsets (SOA for each numeral in the sequence, SOA per numeral). Hence, the SOA per numeral used in the experiments was 33, 67, or 100 ms.

The dark numerals (Fig. 1) on a white background could appear in 12 possible positions in a ring around the continuously presented fixation cross. The eccentricity of the numerals was 1.5° from the viewing distance of 115 cm. The size of each numeral was 0.5° × 0.3° and the line width was 0.05°. In each sequence of numerals, the positions of numerals were chosen randomly with two restrictions to avoid interference between the numerals: in one sequence two or more numerals could not appear in the same position, and two consecutive numerals could not be in adjacent positions. The observer viewed the screen binocularly and was instructed not to move his eyes during the stimulus presentation. At the end of the stimulus presentation, the observer gave his response by typing all of the numerals in the sequence trying to preserve the same order as in the actual sequence. Feedback was given to the observer in two different ways. First, when the observer typed his response, a wrong numeral was indicated immediately by a short tone. Secondly, after the observer had typed the whole sequence, he was shown the correct stimulus sequence at a slower speed. Two highly practiced observers participated in the experiments. One of the observers (J.S.) was the first author.

RESULTS AND DISCUSSION

Before data collection we established that the mask pattern was adequately effective by presenting only one numeral at each trial. This was important for the interpretation of the experimental results because if the stimulus patterns had not

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: SOA, stimulus onset asynchrony.

*Present address: Department of Psychology, University of Helsinki, 00170 Helsinki, Finland.
†To whom reprint requests should be addressed.

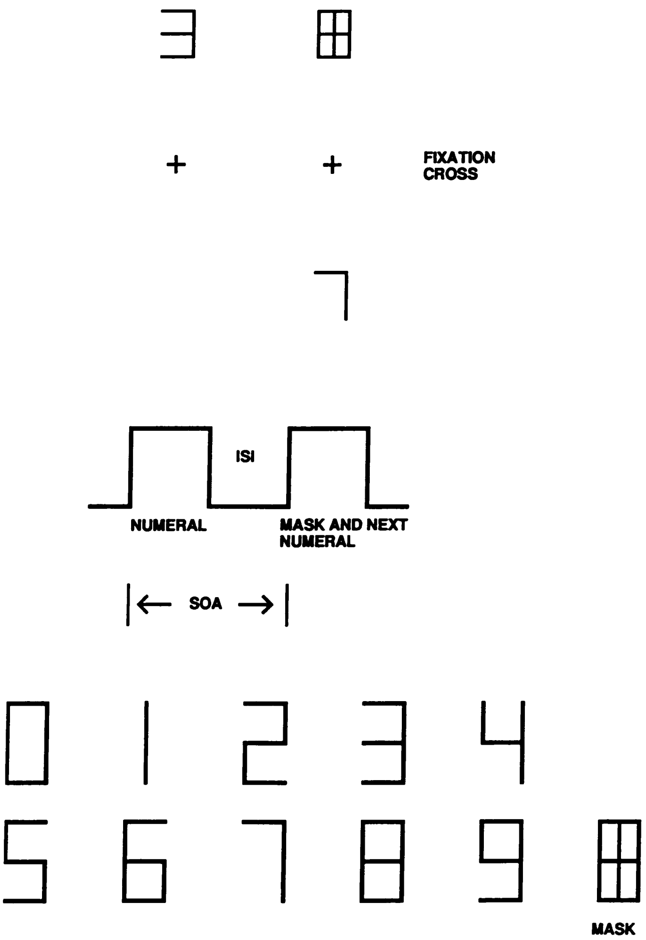

FIG. 1. Temporal presentation of two successive numerals and mask. Below are the shapes of the numerals (0–9) and the mask. ISI (interstimulus interval) is the blank period between the numeral and the mask. SOA (stimulus onset asynchrony) is the time interval between the onsets of the numeral and its mask.

been masked effectively, the observer could have seen the numerals, not one by one as was intended but simultaneously (13). The effectiveness of masking is shown in Fig. 2 (columns for "one numeral").

With increasing number of numerals in the sequence, the proportion of trials in which the observer could recognize all of the numerals in the sequence correctly decreased for all presentation rates (Fig. 2). However, both observers operated significantly above the theoretical chance level even when there were four numerals in the sequence and the SOA per numeral was 67 or 33 ms. Thus, it seems that the speed of attentional shifts is at least as fast as 50 ms per item, a rate suggested by Treisman and Gelade (4) and Bergen and Julesz (5). It has to be noticed that this speed is not necessarily the absolute limit. For instance, the performance might be improved by making the numerals less schematic, thus more recognizable, than they were in the present experiments.

Both observers reported that it was difficult to perceive the temporal order of the numerals in the sequence even though it seemed to be possible to recognize the numerals. This might explain, at least partially, the deterioration in performance with increasing number of numerals in the sequence. To study this introspective report by observers, we analyzed further the results by calculating the proportion of trials where all numerals in the observer's response were correct but not necessarily in the correct order. This kind of plot of the results (Fig. 2) supports the observers' introspective reports: the proportion of trials where all the numerals of the sequence were correct in any order was higher than with preserving order.

In summary, our first finding is that the speed of focal visual attention can be quite fast, at least 50 ms per item, but performance at 33 ms per item is still respectable. Hence, the results support the notion suggested by several authors (4, 5, 8) that the searchlight of attention can be moved four or five times faster than the eyes are able to move. However, the previous evidence for this notion has been based on the indirect inference from experiments, in which the stimulus patterns were displayed simultaneously, and not one by one followed by a mask, as in the present study.

The other finding was that the observers lost information about the order of the numerals in the sequence even though they could still recognize the numerals. This latter finding is in agreement with those of Reeves and Sperling (9), who used in their experiments also the serial presentation of stimulus patterns (but the patterns were always in the same window). It is possible to explain the loss of information about the order of the numerals by simply assuming that after the fast scanning mechanism of focal attention has picked up the

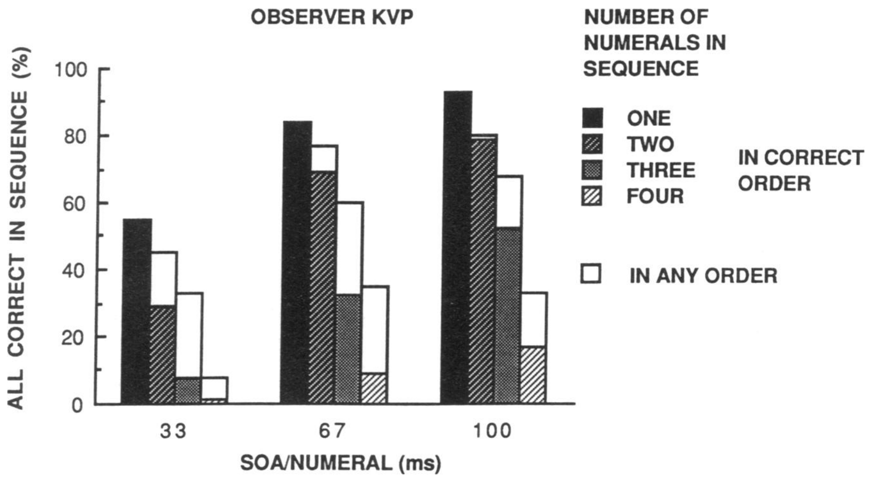

FIG. 2. Percentage of trials where the observer could recognize all of the numerals of the sequence correctly in the correct order (lower part of each column) or in any order (higher, white part of each column) for three SOA per numeral. Each column is an average of four blocks of 30 trials. The theoretical chance levels for recognizing all of the numerals in the sequence in the correct order are 1%, 0.1%, and 0.01% for the sequences of two, three, and four numerals, respectively (the probability of guessing all N numerals of the sequence correctly in the correct order is $1/10^N$). The corresponding chance levels for the responses without preserving the correct order are 2%, 0.6%, and 0.24% (the probability of guessing all N numerals of the sequence correctly in any order is $N!/10^N$).

numerals one by one, guided by the feature gradients in the visual field (6), they enter a visual short-term memory, where the item strength does not depend on the order of the entry [or it is determined by the amount of focal attention that each numeral received (9)], and subsequent reports of the numerals are based on the content of this short-term memory.

We thank Ms. Jih Jie Chang for the complex computer programming and Drs. D. Williams and I. Hadani for the stimulus calibration and for solving a difficult synchronization problem.

1. von Helmholtz, H. (1886) *Handbuch der Physiologischen Optik* (Voss, Hamburg, F.R.G.), Pt. 3, 2nd Ed.
2. Beck, J. & Ambler, B. (1973) *Percept. Psychophys.* **14**, 225–230.
3. Posner, M. I., Snyder, C. R. R. & Davidson, B. J. (1980) *J. Exp. Psychol. Gen.* **109**, 160–174.
4. Treisman, A. M. & Gelade, G. (1980) *Cognit. Psychol.* **12**, 97–136.
5. Bergen, J. R. & Julesz, J. (1983) *Nature (London)* **303**, 696–698.
6. Sagi, D. & Julesz, J. (1985) *Science* **228**, 1217–1219.
7. Sagi, D. & Julesz, J. (1986) *Nature (London)* **321**, 693–695.
8. Efron, R., Yung, E. W. & Nichols, D. R. (1987) *Neuropsychologia* **25**, 637–644.
9. Reeves, A. & Sperling, G. (1986) *Psychol. Rev.* **93**, 180–206.
10. Weichselgartner, E. & Sperling, G. (1987) *Science* **238**, 778–780.
11. Kröse, B. J. A. & Julesz, J. (1989) *Vision Res.* **29**, 1607–1619.
12. Posner, M. I. & Presti, D. E. (1987) *Trends Neurosci.* **10**, 13–17.
13. Hogben, J. H. & Di Lollo, V. (1974) *Vision Res.* **14**, 1059–1069.